BU-72-M

Custodian Copy

PRINCIPLES OF EXPERIMENTATION

Lectures Given

for the

Department of Training Personnel, HSPA

by

Walter T. Federer, Principal Statistician
Dept. of Expt. Statistics, HSPA

and

Consultant in Statistics, PRI

(on sabbatical leave from Cornell University)

* * * * * * *

Monday	Nov. 1, 1954	-	Fundamentals of Scientific Research
Wednesday	Nov. 3, 1954	-	Selection of Treatments in Experimentation and the Factorial Experiments
Thursday	Nov. 4, 1954	-	Experimental and Sampling Designs
Friday	Nov. 5, 1954	-	Analyses of Data
Monday	Nov. 8, 1954	-	Tests of Significance and Statements of Inference

Pencilled notation is for lecture given on Oct 18 1955 to [illegible] group

lecture given to [illegible]

FUNDAMENTALS OF SCIENTIFIC RESEARCH

medical problems

Most research on ~~the growing~~ of sugar cane involves experimentation. and usually [illegible]

humans or animals

Usually the experiment is set up to provide data relating to the tenability of a given hypothesis. Before going further, it might be wise to define the terms that are used in order to be certain that the words used have the same meaning for all of us. Research, in general, refers to the collection and analysis interpretation of data on a basic unsolved problem. The mere collection of facts and figures and the computation of averages is not to be construed as research. Furthermore, re-search should not be confused with research. The nature of a research project is determined by the particular problem under study. To serve more technical aspects, projects are designed to serve purely local or temporary needs. To be a greater scientific value, the research will be on more fundamental problems involving general biological [illegible] laws or principles.

From the data collected on the particular problem, an experimenter wishes to make valid inferences from the data (a sample) to the true and unknown situation (the population from which the sample was drawn), or stated in another way from the particular to the general. This is called the inductive approach. Sir Ronald Fisher (1942) states that inductive inference is the only process known to us by which essentially new knowledge comes into the world. The inductive method is the scientific spirit of the present time. The first requisite of induction is experience to furnish facts. Such experience is obtained by observation and by experiment. The definition of an observation as given by Jevons (1870) is as follows: "To observe is merely to notice events and changes which are produced in the ordinary course of nature, without being able, or at least attempting, to control or vary these

a [illegible]
this category

changes." On the other hand, *experimentation* is a planned and controlled experience. Fisher (1942, page 8) describes experimental observations as "only experience carefully planned in advance, and designed to form a secure basis of new knowledge; that is, they are systematically related to the body of knowledge already acquired, and results are deliberately observed, and put on record accurately." The importance of the experiment is well summarized by Jevons (1870); thus: "It is obvious that experiment is the most potent and direct mode of obtaining facts where it can be applied. We might have to wait years or centuries to meet accidentally with facts which we can readily produce at any moment in a laboratory ...."

Empirical laws

Probably the best way to explain the term *hypothesis* is to relate this term to two other words -- *theory* and *law*. The difference between the hypothesis, the theory, and the law is in the degree of surety of the absolute. When an idea is suggested by an observed phenomenon, it is spoken of as a hypothesis. It represents a desire to explain the phenomenon such as, for example, a method by which plants take food from the soil. The hypothesis is important in the deductive method in that certain facts are assumed to be true and then from these facts the statement in the hypothesis is deduced. The truth of a hypothesis depends upon subsequent verification. A theory is a limited and inadequate verification of a hypothesis, while a law is a complete verification of a hypothesis. An example of where a hypothesis became a theory and a theory became a law is represented by the work of Mendel. Mendel hypothesized that the genetic segregation in segregating generations follows definite ratios. A theory was established from his work on numerous characteristics on peas. Since the theory on Mendellion segregation has been verified beyond all reasonable doubt, it has become a law.

There are certain advantages to the formulation of hypotheses:

(i) ~~it~~ the hypothesis correlates facts,
(ii) it forecasts other facts, and
(iii) it allows for discrimination between valuable and useless information.

Every experiment is the result of a tentative hypothesis thought out in advance of the actual test. The hypothesis is based on the recognition of coincident phenomena, or upon a familiarity with possible causes and effects. A good hypothesis should possess the following qualities

(i) it should be plausible,
(ii) it should be capable of proof, that is, it should provide a susceptible means to attack the problem created thereby,
(iii) it must be adequate to explain the phenomena to which it is applied,
(iv) it should involve no contradiction,
(v) a simple hypothesis is preferable to a complex one.

There is little use in formulating a hypothesis on a complex basis unless it is possible to collect the data on which it may be verified. Formulation of several hypotheses may be useful, even though some of the hypotheses are wrong, in order to eliminate particular ideas from the problem. At any time, the investigator must be ready to abandon a hypothesis or a theory when further data proves that previous views are untenable.

The principles involved in scientific experimentation have been stated in many ways and in many places. A number of books have been written on this subject. The books by Wilson (1952) and Churchman (1948) are recommended for reading since they include a discussion of scientific experimentation in light of present statistical knowledge. Also, articles on experimental methods, on statistical analyses, and on the interpretation of results appear in a number of scientific journals. ~~Leonard and Clark (1939) have enunciated these principles for agronomists~~. The principles of scientific experimentation are listed below.

1) Formulation of questions to be answered and hypotheses to be tested.

The scientific experiment should be set up to answer a specific question or questions. Precise formulation of the question (or questions) to be answered enables the experimenter to state his hypothesis in more precise terms and to plan his experimental procedure more effectively. Clear and precise questions and hypotheses at this stage enable the experimenter to proceed rapidly to the next step. It is advisable to have a statement of the questions and hypotheses in written form.

2) A critical and logical analysis of theproblem raised.

After formulation, the hypothesis should be critically and logically evaluated. A review of pertinent literature is a valuable aid in evaluating a hypothesis. The reasonableness and the utility of the aims of the experiment should be carefully considered. After a critical evaluation of the hypotheses in step 1, the experimenter may find it advisable to reformulate the questions and hypothesis before proceeding with the experiment. Also, the experimenter shoulld evaluate the possible outcomes of the experiment in terms of relative cost and economic returns from such results.

3) Selection of a procedure for research.

The experimental procedure to be followed depends to a large extent upon the field in which the research is being conducted. The selected experimental procedure in most fields of research will involve some or all of the following considerations:

(i) The selection of treatments ~~or~~ entries to be included in the experiment,
(ii) the selection of characteristics to be measured,
(iii) the selection of the unit of observation, number of replications and the sampling or experimental design (the latter should be amenable to statistical analysis),
(iv) the control of the effect of adjacent units on each other or border effect,

(v) an outline of pertinent summary tables and probable results,
(vi) an outline of analyses to be performed,
(vii) a statement of cost in terms of material, personnel, equipment and so forth.

After considering the above items, the experimenter may decide to reformulate the hypothesis, to change the experiment, or to continue to the next step in scientific experimentation. The formulation of hypotheses and the selection of treatments are extremely important and closely associated considerations. A large part of the success of an experiment may depend upon the correct selection of treatments, Also, the selection of an appropriate experimental or sampling design is of considerable importance in the testing of a hypothesis and in estimating treatment effects. [illegible] (i) [illegible] (ii) to [illegible] with minimum expenditures.

4) Selection of suitable measuring instruments.

In any experiment, the experimenter must select measuring instruments that are sufficiently accurate for the purposes at hand. He should be constantly on guard against sources which might add to the variability of his experimental material. Remember that no chain is stronger than its weakest ~~length~~ link. Some suggestions of items to guard against are:

(i) the scales on which experimental material is weighed should be checked periodically,
(ii) the sampling procedure should be sufficiently precise in order that the variability is not unduly influenced by the sampling procedure.

[illegible]

For example, tonnage of cane per acre may be precise enough but the methods of collecting the juice samples may allow considerable variability in the results. This affects the tons of cane per acre as well as analyses on juices.

(iii) the management of the experimental ~~plots within a block~~ units in the [illegible] of the expt should be uniform. (~~Particular care should be given~~ to ~~irrigation and other cultural practices~~).

(iv) Enough significant figures (see Snedecor, 1946, chapter 5 and Yates, 1937, page 91) should be carried in the analysis to be certain that this does not increase the variability in the experimental material.

5) Control of the personal equation.

The experimenter must select an experimental procedure that is free from biases or favoritisms. In some studies, it is the practice to observe 3 samples and to discard the most discrepant one. The procedure goes even further in the so-called "intelligent placement" of treatments where the favorites may be placed under the best conditions, and if a "favorite" is lower than preconceived, the results are discarded. If the experimenter follows the procedure of taking the two samples which agree best and discarding the third sample, which is the most divergent one, he should not use ordinary statistical procedures for summarizing his results. Procedures have been developed for this sampling method. Also, we could use statistical procedures involving rank order statistics which would ~~take care of~~ does not give undue wt. to a discrepant observation

The question of discarding units of observations often arises in experiments in which the units of observations are not of equal value in the contribution to the treatment mean. For example, a ~~field plot or~~ laboratory sample for one of the treatments may be destroyed or damaged due to the carelessness of the technician; ~~a group of the experimental plants may become diseas~~ed, ~~adverse weather conditions or rats may destroy or damage part of the experiment; and so fo~~rth. Now, the question may arise as to whether or not zeros or abnormally low values from the damaged experimental units should be retained in ~~his~~ the statistical analysis or whether they should be deleted prior to the statistical analysis. In this regard, experimenters are urged to consider the following rule. <u>If the observation contributes to the true treatment difference, it should always be retained. If the observation does not contribute to the true treatment difference, it should be discarded,</u>

provided the reasons for discarding the observation is one that would be held valid by the experts in that field.

The problem of controlling personal biases or favoritisms usually does not require elaborate precautions. ~~Often simple procedures suffice~~. For example, in the analyses of samples in the laboratory, a random assignment of numbers to the samples rather than identification of the sample often provides unbiased results which are not possible when the technician knows the identity of the samples. Al~~so, it is a good idea to number the field plots consecutively and to identify them in no other way~~. The treatment designation should not be included on the tag identifying the ~~plot~~ sample. Also, the notebook should not contain the identification of the treatment. If the identifications are not present, then there is a good possibility that the person doing the grading, ~~or~~ judging, laboratory worker ~~of the plot~~ will give it an unbiased score.

6) A complete analysis of the data and interpretation of results in light of experimental conditions and hypothesis tested.

Statistical procedures are valuable aids in reducing the data to summary form. Remember that statistical methods represent a kit of tools [surgical instruments] and not a set of crutches. The results of an experiment must be interpreted in light of the statistical evidence obtained and the theoretical considerations of the subject under experimentation. The "interpretation" of the results of an experiment does not end with calculations of a mean, of an F value, of a confidence interval, or of a LSD! In testing hypotheses and in estimating effects of the treatments, it is essential to use the appropriate statistical procedures. The resulting computations should be checked to ensure against computational errors.

7) Preparation of a complete, correct, and readable report of the experiment.

The pertinent results of an experiment should be completely and carefully reported in written form. One should present enough of the summary data to allow others to test the various hypothesis. By giving an account of the experimental procedure and statistical methods (or references to statistical methods) used the reader may decide for himself whether or not he considers the results sound. Always keep this in mind, an experiment is not conducted to show "significant results". It is conducted to give results pertaining to the questions raised and the hypothesis stated. The true scientist does not set out to prove a point but rather to accumulate evidence on the problem at hand.

In the above seven steps, the various principles of scientific experimentation have been outlined. It should be emphasized that the above principles are not entirely statistical, nor does every step involve statistics; perhaps here would be a good place to define what is meant by statistics. Snedecor (1946, chapter 1) defines statistics in the following way; it consists of the three things:

(i) Designing experimental procedures or sample surveys,
(ii) collection and summarization of the data derived from the experiment or the survey,
(iii) making inferences from the facts in the sample to the true population from which the sample was drawn.

The first step is misused continually. People use poorly designed experiments and surveys. A small amount of reading could alleviate this difficulty for most people. The collection of the data often gives rise to considerable difficulty. This is especially true of sample surveys. The third step seems to cause considerable trouble -- the drawing of inferences. In our everyday experiences, we are continually making inferences. We have a set of facts before us and we make certain decisions based upon certain risk elements. For example, we look at the paper and out-of-doors, and we infer whether or not it will rain. This inference and a consideration of the risks involved will lead us to the decision of taking our raincoat or of not taking our raincoat. We see facts and figures in an advertisement of a local newspaper. From these facts and figures, we draw certain inferences, either the product is good or not good, or maybe mediocre. Try as we might, we just cannot get away from the third step. And remember, this is the difficult step. The second step can be put on a purely routine or cook book nature. We can give you a recipe for doing various computations, but it is the designing of experiments and sample surveys and the drawing of inferences which is the difficult part of statistics. Due to the fact that we conduct experiments and that we do draw inferences, we just cannot escape using statistics no matter what the field of experimental research.

Principles 1 and 2 are entirely statistical in nature and are completely within the scope of the field in which the experiment is performed. In principle no. 3, steps i, ii, iv, vii, are non-statistical, v may or may not be statistical, and steps iii, vi are entirely statistical in nature. Principles 4 and 5 are partly statistical, and principle 6 is highly statistical while principle 7 is non-statistical. Hence, the design of an experiment is just one item to be considered in planning and conducting an experiment and in interpreting and reporting the experimental results.

In the above, only the high spots in experimentation have been touched upon. It is suggested that the reader refer to such works as Churchman (1948), ~~Cochran and Cox (1950, chapter 1)~~, Cohen and Nagel (1934), ~~Federer (1955, chapter I)~~, Jevons (1870), ~~Leonard and Clark (1939, chapters 1-4~~ and chapter 20), Love ~~(1943, pages 155-169, 183-190 and 197-219)~~, Wilson (1952), Bross + Huff ~~Wishart (1940), Wishart and Sanders (1936)~~. A number of these references are philosphical in nature but others Bross & Huff amusing reading / represent ~~practical appli-~~ cations of these results. The work of Wishart and Sanders has been used by experimenters for a considerable time. Also, the book by Leonard and Clark represents easy reading.

Statisticians in medical work
Berkson at Mayo Clinic
[illegible] at [illegible]
Lancaster at Melbourne Univ [illegible]
Cornfield [illegible]
Bross [illegible]

W. T. Federer
Principal Statistician
Dept. of Expt. Statistics, HSPA

10/27/54
mm

Revised 4/13/55
mm